

DOCUMENT RESUME

ED 208 031

TM 810 726

AUTHOR Krus, Patricia; Maher, John  
TITLE Assessing Teacher Skills Needed to Implement Objectives-Based Instruction.  
PUB DATE [79]  
NOTE 15p.  
EDRS PRICE MF01/PC01 Plus Postage.  
DESCRIPTORS Course Objectives; \*Criterion Referenced Tests; Education Majors; Elementary Secondary Education; Higher Education; Preservice Teacher Education; \*Teaching Skills; \*Test Construction

ABSTRACT

The purpose of this paper is to describe the design and construction of three objectives-based tests to assess preservice teachers' performance in an instructional skills course. The three tests covered: writing worthwhile objectives, designing effective instruction, and assessing student learning. All tests were analyzed for item difficulty, item discrimination, and reliability. With few exceptions, performance on each skill was significantly correlated with total test performance, but not necessarily with other skills.  
(BW)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

"PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

J. Maher

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

## Assessing Teacher Skills Needed to Implement Objectives-Based Instruction

Patricia Krus and John Maher

Arizona State University

U.S. DEPARTMENT OF EDUCATION  
NATIONAL INSTITUTE OF EDUCATION  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

✓ This document has been reproduced as  
received from the person or organization  
originating it.

□ Minor changes have been made to improve  
reproduction quality.

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official NIE  
position or policy.

ED208031

The comparison of learners' performance to the performance specified in objectives is an important characteristic of objectives-based instructional programs. Criterion-referenced tests comparing learners' performance to achievement of learning objectives rather than to the performance of other learners were introduced by Glaser (1963) and Popham and Husek (1969). It has been asserted (Swaminathan, Hambleton & Algina, 1974) and demonstrated (Gibbons, 1970; Heathers, 1972) that the successful implementation of an objectives-based instructional program depends on efficient assessment.

TM 810 726

The usual procedure for objectives-based test development is the determination of a well-defined behavior domain (Popham, 1975), the keying of the behavior domain to objectives, and construction of test items to assess learner achievement for each program objective. Of critical importance is the correspondence of the learning objective and test items with respect to the performance called for, the conditions under which the performance occurs, the criteria by which the performance is assessed, and the range of content assessed (Higgins & Sullivan, 1978).

The purpose of this paper is to describe the design and construction of three objectives-based tests to assess preservice teachers' performance in an instructional skills course. The validity, reliability, and practicality of the tests will be

addressed. The instructional skills assessed include writing worthwhile objectives, designing effective instruction, and assessing student learning. This paper is the second of three papers detailing the development and evaluation of the course.

### Method

Subjects. The sample consisted of 74 preservice teachers in the College of Education, Arizona State University, Tempe, Arizona. Included in the sample were only those students who completed the course and for whom measurements on all the unit tests were available

Materials. Three instructional skills tests corresponding to the three units of instruction were constructed. Sample tests are available from the Department of Educational Technology, Box FLS, Arizona State University, Tempe, AZ 85281. Items on the three tests sampled eleven subject areas and representative grade levels from kindergarten through twelfth grade. Items also sampled three learning domains (cognitive, affective, psychomotor). The items were produced by different combinations of curricular area X grade level X learning domain. Items were selected from the pool in accordance with the strata already mentioned for each of the following instructional skills:

Unit 1. Writing Worthwhile Objectives

- a. Distinguishing objectives from activities
- b. Identifying worthwhile objectives
- c. Identifying well-written objectives
- d. Writing objectives

Unit 2. Designing Effective Instruction

- a. Identifying instructional information
- b. Providing instructional information
- c. Identifying appropriate practice
- d. Providing appropriate practice activities
- e. Identifying effective practice activities
- f. Identifying effective feedback activities

- Unit 3. Assessing Student Learning
- a. Identifying appropriate assessment items
  - b. Identifying well-written test items
  - c. Writing assessment items

Five to twenty items were written to assess each instructional skill. The number of items for assessing each instructional skill was determined by the complexity of the skill. Ten to twenty items were used for each skill that involved identification (selected responses) and five items were used for each skill that involved writing (constructed response).

In constructing the test, there was an attempt to achieve a balance of correct and incorrect items used to assess each skill. The test items assessed each of the principles (rules) taught in the course. When complex principles were included, that is, when a principle involved two or more concepts, each concept was tested individually and in combination with the other concepts in the principle. For example, in assessing the skill of identifying well-written test items, teachers were taught that a test item must be (1) clear, and (2) free of prompts. Prompting can be further broken down into (1) prompting with specific determiners (T-F items), (2) prompting with implausible alternatives (Multiple Choice items), (3) prompting with equal choices (Matching items), and (4) prompting with grammar, plural words, and so forth. In constructing items to assess teacher performance on this skill, each of these characteristics was tested individually and in combination with the others. Constructing the items in this manner ensured the content validity of the tests.

Efforts were made to assure the practicality of test administration for large enrollment multi-section courses. Tests were constructed to be easily completed in a 75-minute class period (typically our classes meet once or twice per week). In addition, each test can be split into subtests for administration over two class periods. The time to score each test typically ranged from three to five minutes. Tests from a class of 30 teachers can be scored in 1.5 to 2.0 hours.

Procedures. The three unit tests--Writing Worthwhile Objectives, Designing Effective Instruction, and Assessing Student Learning--were randomly distributed to students on the first class day to determine prior knowledge of the course material. As students completed each unit, they were given a posttest on that unit. Data were collected on performance by objective and performance by unit test. All tests were analyzed for item difficulty, item discrimination, and reliability.

### Results and Discussion

Item Difficulty and Discrimination. Item analysis results for pretest and posttest by individual unit are shown in Table 1. The average pre- post- item discrimination on the three tests was approximately .27. Item difficulty indices increased from pretest to posttest on all three tests, indicating that test items became easier with instruction.

Performance in Selected-Response and Constructed-Response Skills. With one exception, student performance was higher on skills involving selected responses (identifying) than those involving constructed

responses (writing) (See Table 2). This indicates that tasks requiring selected responses were easier for the students than those requiring constructed responses. The exception to this trend occurred in Writing Worthwhile Objectives, where students performed better on the task of writing objectives than on the task of identifying well-written objectives. This discrepancy might be due to the minute discrimination required for identifying well-written objectives in contrast to the more subjective criteria used for assessing student performance or writing objectives.

Test Reliability. Internal consistency reliability was measured by Kuder-Richardson's Formula 20. Reliability coefficients for the posttests for each unit were moderate, ranging from .55 to approximately .66. The restricted variances of the posttests of the instructed group were expected due to the nature of skills-based instruction. When the three unit tests were given to an uninstructed group of students, greater variance in scores was obtained as were higher reliability coefficients (see Table 3). The restricted variance for the Assessing Student Learning test for the uninstructed group contributed to a lower reliability coefficient than for the other two tests. This trend for Assessing Student Learning was repeated in the results of the test with the instructed group.

Intercorrelations of Skills Within a Unit. Posttest intercorrelations of skills within a unit ranged from  $-.11$  to  $.78$ , with the majority (68%) being statistically significant. In Writing Worthwhile Objectives, pupil performance on each skill correlated

significantly with total test performance. In addition, performance on the skill of distinguishing objectives from activities was significantly correlated with performance on each of the other skills. Performance on the skill of identifying worthwhile objectives was not significantly correlated with performance on writing objectives (see Table 4). It would appear that identifying well-written objectives and writing objectives are discrete skills.

The intercorrelations for skill within unit for Designing Effective Instruction are shown in Table 5. While performance on each of the skills in this unit is significantly correlated with performance on the total test, there are few intercorrelations among the skills themselves. Performance on the skill of identifying instructional information was significantly correlated with performance on both identifying ( $r = .31$ ) and providing appropriate practice ( $r = .22$ ). Performance on the skill of providing appropriate practice was also significantly correlated with performance on providing instructional information ( $r = .34$ ) and identifying appropriate practice ( $r = .20$ ). Performance on providing information correlated significantly ( $r = .19$ ) with performance on identifying effective practice activities. Finally, performance on identifying effective practice activities correlated significantly ( $r = .60$ ) with performance on identifying effective feedback activities.

The intercorrelations for skill within unit for Assessing Student Learning are shown in Table 6. As in the first two units, performance on each of the skills taught correlated significantly with the total test. Performance on the skill of writing assessment items was

significantly correlated with performance on both identifying appropriate items ( $r = .21$ ) and identifying well-written items ( $r = .21$ ), a relationship that was expected.

It is interesting to note that, with few exceptions, pupil performance on each skill was significantly correlated with total test performance, but not necessarily with other skills. Generally, even when the skills are significantly correlated, the correlations are low, with only six of the twelve significant intercorrelations among skills above .30. This trend tends to indicate that, overall, the skills in each unit are relatively discrete. As a result, we think each of the skills in the course should be taught. The assumption that performance on one skill would transfer to other skills, even within an individual unit, would be invalid.

## References

- Heathers, G. Overview of innovations in organization for learning. Interchange, 1972, 3, 47-68.
- Higgins, N. & Sullivan, H. Writing Worthwhile Objectives, POB 27553, Tempe, AZ: Teaching for Competency, 1978.
- Popham, W.J. Education evaluation. Englewood Cliffs, NJ: Prentice-Hall, 1975.
- Popham, W.J., & Husek, T.R. Implications of criterion-referenced measurement. Journal of Educational Measurement, 1969, 6, 1-9.
- Swaminathan, H., Hambleton, R.K. & Algina, J. A decision-theoretic approach to issues in criterion-referenced assessment. Laboratory of Psychometric and Evaluative Research Report No. 1. Amherst, Mass.: School of Education, University of Massachusetts, 1973.

Table 1  
Mean Item Difficulty and Discrimination for  
Pretest and Posttest by Unit

| Unit                            | Pretest | Posttest |
|---------------------------------|---------|----------|
| Writing Worthwhile Objectives   |         |          |
| Difficulty                      | .56     | .88      |
| Discrimination                  | .32     | .24      |
| Designing Effective Instruction |         |          |
| Difficulty                      | .47     | .88      |
| Discrimination                  | .30     | .24      |
| Assessing Student Learning      |         |          |
| Difficulty                      | .87     | .89      |
| Discrimination                  | .31     | .23      |

Table 2  
Mean Pretest-Posttest Percentage Scores

| <u>Writing Worthwhile Objectives</u>   | <u>Pre</u> | <u>Post</u> |
|----------------------------------------|------------|-------------|
| Distinguish Objectives from Activities | 75         | 90          |
| Identifying Worthwhile Objectives      | 64         | 91          |
| Identifying Well-Written Objectives    | 61         | 83          |
| Writing Objectives                     | 64         | 90          |
| <u>Designing Effective Instruction</u> |            |             |
| Identifying Information                | 68         | 92          |
| Writing Information                    | 45         | 85          |
| Identifying Appropriate Practice       | 46         | 94          |
| Describing Appropriate Practice        | 43         | 89          |
| Identifying Effective Practice         | 54         | 94          |
| Identifying Feedback Procedures        | 46         | 92          |
| <u>Assessing Student Learning</u>      |            |             |
| Identifying Appropriate Items          | 71         | 90          |
| Identifying Well-Written Items         | 64         | 89          |
| Writing Assessment Items               | 53         | 86          |

Table 3

Test Internal Consistency Reliability for Uninstructed  
and Instructed Groups

| Test                            | Uninstructed<br>Group | Instructed<br>Group |
|---------------------------------|-----------------------|---------------------|
| Writing Worthwhile Objectives   | .80                   | .63                 |
| Designing Effective Instruction | .77                   | .66                 |
| Assessing Student Learning      | .63                   | .55                 |

Table 4  
Intercorrelations of Skills: Objectives Unit

| Skill                                  | Skill |       |       |       | Total<br>Test |
|----------------------------------------|-------|-------|-------|-------|---------------|
|                                        | 1     | 2     | 3     | 4     |               |
| 1. Distinguish objectives/activities   | 1.0   | .32** | .33** | .22** | .56***        |
| 2. Identifying worthwhile objectives   |       | 1.0   | .14   | .42** | .66***        |
| 3. Identifying well-written objectives |       |       | 1.0   | .04   | .48***        |
| 4. Write objectives                    |       |       |       | 1.0   | .68***        |

\*  $p < .05$   
 \*\*  $p < .01$   
 \*\*\*  $p < .001$ .

Table 5  
Intercorrelations of Skills: Instruction Unit

| Skill                                    | Skill |     |       |        |      |        | Total<br>Test |
|------------------------------------------|-------|-----|-------|--------|------|--------|---------------|
|                                          | 1     | 2   | 3     | 4      | 5    | 6      |               |
| 1. Identifying instructional information | 1.0   | .16 | .31** | .22**  | .04  | -.01   | .35***        |
| 2. Provide information                   |       | 1.0 | -.11  | .34*** | .19* | .03    | .43***        |
| 3. Identifying appropriate practice      |       |     | 1.0   | .20*   | .11  | .13    | .40***        |
| 4. Provide practice                      |       |     |       | 1.0    | .11  | .04    | .45***        |
| 5. Identifying effective practice        |       |     |       |        | 1.0  | .60*** | .73***        |
| 6. Identifying appropriate feedback      |       |     |       |        |      | 1.0    | .60***        |

\*p<.05

\*\*p<.01

\*\*\*p<.001

Table 6  
Intercorrelations of Skills: Assessment Unit

| Skill                                        | 1   | Skill<br>2 | 3    | Total<br>Test |
|----------------------------------------------|-----|------------|------|---------------|
| 1. Identifying<br>appropriate<br>test items  | 1.0 | .00        | .21* | .44***        |
| 2. Identifying<br>well-written<br>test items |     | 1.0        | .21* | .69***        |
| 3. Writing<br>test items                     |     |            | 1.0  | .78***        |

\*p<.05

\*\*p<.01

\*\*\*p<.001